The AAPS Journal 2007; 9 (2) Article 13 (http://www.aapsj.org).

Themed Issue: Bioanalytical Method Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays
Guest Editors - Mario L. Rocci Jr., Vinod P. Shah, Mark J. Rose, Jeffrey M. Sailstad

Best Practices During Bioanalytical Method Validation for the Characterization of Assay Reagents and the Evaluation of Analyte Stability in Assay Standards, Quality Controls, and Study Samples

Submitted: February 7, 2007; Accepted: April 2, 2007; Published: April 20, 2007

William Nowatzke[1] and Eric Woolf[2]

[1]Amgen, Inc, Department of Pharmacokinetics and Drug Metabolism, Thousand Oaks, CA 91320
[2]Merck Research Laboratories, Department of Clinical Drug Metabolism, West Point, PA 19486

Abstract

Characterization of the stability of analytes in biological samples collected during clinical studies together with that of critical assay reagents, including analyte stock solutions, is recognized as an important component of bioanalytical assay validation. Deficiencies in these areas often come to light during regulatory inspections. Best practices, based on current regulatory guidance, for the assessment of these issues as they pertain to ligand binding and chromatographic assays are covered in this review. Additionally, consensus recommendations reached during the recent AAPS/FDA Workshop on bioanalytical assay validation are highlighted.

Keywords: Stability, biological samples, test article, bioanalytical, ligand binding assay, chromatographic assay

Introduction

The successful development of novel pharmaceuticals cannot be achieved without the use of data generated using validated bioanalytical methods. Several key documents,[1] journal publications,[2-4] and scientific meeting presentations[5] are available to guide ligand binding and chromatographic bioanalytical laboratories in the development, validation, and use of bioanalytical methods.

The generation of data pertaining to analyte and reagent stability is a key requirement of the assay development/validation guidelines. Specifically, data pertaining to the stability of (1) analyte stock solutions and chemical reagents used for bioanalytical methods, (2) biological fluid samples containing analytes of interest, and (3) extracts of biological fluid samples are required to be generated during the assay validation process. Descriptions of methods to assess these criteria are described below.

Corresponding Author: William Nowatzke, Amgen, Inc, Department of Pharmacokinetics and Drug Metabolism, One Amgen Center Drive, MS 30E-3-B, Thousand Oaks, CA 91320. Tel: (805) 447-2151; Fax: (805) 376-2069; E-mail: nowatzke@amgen.com

Standard Reference Material

The vast majority of bioanalytical assays, regardless of the method of analysis, use calibration standards prepared by spiking control matrix with a solution containing the analyte(s) of interest. Hence, an adequately characterized standard reference material must be available from which these solutions may be prepared. Characterization, with respect to identity and purity, of the reference standard is often conducted by quality control (QC) groups associated with the manufacturing process. Documentation of the characterization must be available to the bioanalytical laboratory when this material is used for method validation and sample analysis. Characterization should include assessments of the compound identity, the purity, the concentration or titer of the material, the storage conditions, and the expiration date. If possible, the reference standard should be identical to the analyte. When this is not possible, an established chemical form (free base or acid, salt or ester) of known purity can be used.[1] In the case of biological analytes (eg, antibodies or peptides/proteins) it is desirable that the reference standard used for the characterization of stability be the exact formulation administered to test animals or clinic study participants.

Expiration/retest dates are typically assigned to reference batches of the neat analyte. The consensus reached during the AAPS/US Food and Drug Administration (FDA) workshop was that these expiration dates do not automatically pertain to solutions prepared from a given batch of neat reference material[4]; the stability of solutions prepared using the material needs to be assessed separately (see below).

Chromatographic-based assays often use internal standards. Compounds used as internal standards are typically prepared in small (<1 g) batches and are hence not available in the quantities necessary for full analytical characterization. In recognition of this point, a recommendation of the AAPS/FDA workshop was that a certificate of analysis should not be required for materials used as internal standards in bioanalytical methods. Rather, only the lack of interference between analyte and internal standard needs to be demonstrated during assay validation. For a chromatographic-based assay, this assessment may be conducted by injecting a mass

of internal standard (the equivalent of that anticipated to be present in samples) into the chromatographic system and demonstrating that any signal present in the detection channel of the analyte at the analyte retention time will not affect analyte quantitation under assay conditions.

Formulation Stability

While typically assessed outside of bioanalytical laboratories, the stability of the analyte in the dosage form in which it is administered (ie, the formulation) must be demonstrated to ensure the validity of the conclusions that are based on the bioanalytical data generated from a study in which a particular formulation is administered to test animals or humans. Demonstration of the stability of the manufactured formulation of the test article is mandated by all regulatory agencies. For example, in the United States, the FDA Good Laboratory Practice (GLP) Regulations, 21 CFR Part 58, Section 58.105(b), states: "The stability of each test or control article shall be determined by the testing facility or by the sponsor either: (1) Before the initiation, or (2) concomitantly according to written standard operation procedures, which provide for periodic analysis of each batch."

Similar requirements exist for the Japanese Ministry of Health, Labor, and Welfare, GLP Standards, Ordinance 21; and the Organization for Economic Cooperation and Development Principles of Good Laboratory Practice, C(97) 186/Final.

As was the case for reference standards, characterization of formulated test articles is often conducted outside of the bioanalytical laboratory. This being the case, a presentation of methods to assess formulation stability is outside of the scope of this review. Nevertheless, bioanalytical laboratories should ensure that they have access to formulation stability information in the event that these data are requested during a quality audit.

Study Sample Stability

For the purposes of this discussion, study samples are defined as biological samples collected following the administration of a test compound to animals or humans. Evaluation of the stability of targeted analytes in the test biological matrix is critical to validating a bioanalytical method. The integrity of study sample data can be ensured only if supporting stability data are available to confirm that degradation after sample collection has not occurred.

Ideally, study sample stability would be evaluated using freshly collected matrix from a test subject, and processing and storage of the sample would occur in the same way it does for all study samples. Upon collection, stability evaluation would continue over a period of time sufficient to cover the anticipated length of time between sample collection and analysis. This approach is often impractical, as stability testing should be completed, or should be well in progress, prior to study initiation. In addition, an accurate baseline value cannot be defined, as some degree of metabolism and excretion can be expected to have occurred in the test subject.

In lieu of a fresh collection scenario, stability samples may be prepared by adding the reference standard of the analyte to control biological matrices that are representative of those anticipated to be collected during the study. The consensus reached during the AAPS/FDA workshop was that the use of a "stripped" or otherwise altered matrix for study sample stability evaluation is unacceptable[4]; the matrix used for sample stability experiments should mimic that anticipated to be received for analysis. Note that if the study sampling protocol specifies that stabilizers or other additives are to be added to the samples at the time of collection, it is appropriate to use matrices treated in such a manner for the stability assessment experiments.

The matrix should resemble that of the test system as closely as possible. In ideal situations, a predose biological matrix from the animal species to be studied is used. More commonly, a pool of matrix will be prepared by mixing predose biological matrix samples from several different animals or subjects. The pooled matrix will then be used for the preparation of the stability samples needed to perform the experiments described below.

In the case of ligand binding assays, it may be appropriate to screen the pooled matrix prior to the preparation of stability samples to ensure adequate assay response. Adequate assay response should be obtained from samples prepared using the pooled matrix at least at the lower limit of quantitation (LLOQ) and at the low-QC concentration.[3]

Study sample stability should be assessed in each matrix (ie, plasma, urine, etc) in which analyte is to be quantitated. Furthermore, it is generally recognized that stability assessments are species-specific. For example, the fact that an analyte's stability has been demonstrated in dog plasma does not ensure similar stability in dog urine (same species, different matrix) or in human plasma (same matrix, different species).

For plasma samples in particular, the anticoagulant used during sample collection should match that used for the collection of control matrix. The current consensus is that a change of anticoagulant anion necessitates a reassessment of stability.[4] For example, stability of the analyte in heparinized plasma does not guarantee stability in ethylenediamine-added (EDTA-added) plasma. Conversely, no consensus exists as to whether a change in anticoagulant counter ion (sodium heparin to lithium heparin or sodium EDTA to potassium EDTA) necessitates such a reassessment.[4]

Current guidance[1] requires the assessment of long-term study sample storage stability, study sample freeze-thaw stability,

and study sample bench-top stability. These assessments are best performed by preparing sets of stability samples at a minimum of 2 concentrations (3× the assay LLOQ and 0.75× the assay upper limit of quantitation [ULOQ]). These samples are typically prepared by adding small volumes of concentrated stock solutions of the analyte to a volume of the control matrix. To approximate study samples as closely as possible, it is recommended that the volume of stock solution used for the preparation of the stability samples be no more than 1% of the volume of control matrix. For example, no more than 1 mL of stock solution should be used for the preparation of a 100-mL stability sample pool. It is recognized that circumstances (eg, limited matrix, availability of concentrated stock solutions, accuracy of pipetting small volumes) may prohibit adherence to this recommendation.

Following preparation, the stability sample pools are aliquoted into individual sample tubes similar to those in which it is anticipated that study samples will be received. Ideally, the material from which the stability sample storage tubes are manufactured should be identical to that in which the study samples will be received. For example, just because the integrity of samples stored in polypropylene has been demonstrated does not necessarily mean that samples stored in tubes made of a different material (eg, glass) will be stable. The individual aliquots are then stored under the conditions that will be used for study samples. No clear consensus exists as to whether demonstration of analyte stability in samples stored at 1 temperature (eg, –20°C) automatically implies stability at a lower temperature (eg, –70°C).[4] Appropriate documentation of the preparation and storage of the QC samples that are intended to be used to establish study sample stability is paramount.

Long-Term Analyte Stability

Long-term storage stability assessment experiments should be designed to confirm analyte stability in the test system matrix covering the length of time from sample collection to sample analysis. Such an assessment gives credibility to the final study data. Long-term storage stability is assessed through the repeated analysis of stability sample aliquots, prepared as described above, over the assessment time frame.

It is critical that quantitation of stability samples be made against freshly spiked standards. The matrix standards against which frozen and then thawed samples are quantitated should not have been previously frozen unless stability under those storage conditions has been demonstrated.[3]

Typically, replicate frozen aliquots of the stability samples are assayed and the mean results are used to evaluate stability. The period of time allowed to elapse between assessments may be variable. The first few assessments are usually made on a daily basis. Once stability has been demonstrated during these initial assessments, the period of time between assessments may be extended to weekly, then monthly or even less frequently. Such assessments are continued until the maximum period of time anticipated to elapse between sample collection and sample analysis is attained.

Criteria for long-term stability assessment are not specifically defined in currently available guidance documents; it is simply stated that nominal concentrations should be used.[1] At least 1 publication[3] has recommended that stability be evaluated using the same acceptance criteria applied to evaluate the accuracy and precision of QC samples.

One experimental approach for assessing long-term stability proceeds as follows. Stability samples are prepared as described above. Once prepared, replicate aliquots of the stability samples are analyzed against freshly prepared matrix standards. The initial (day 0 or day 1) analysis should take place within 24 hours of stability sample preparation. Subjecting the stability samples to a freeze-thaw cycle prior to their initial analysis is optional. The results of the initial analysis of the stability samples are used primarily to assess the accuracy of the preparation of the stability samples as well as the short-term stability of the analyte in the matrix.

The guidance[1] specifies that nominal concentrations should be used to assess stability. It is recommended that the measured mean concentrations derived from the initial analysis of the stability samples be within ~5% to 7% of the nominal concentrations. There may be circumstances (eg, the use of incurred study samples) where baseline target values are defined other than by nominal concentrations. In such cases the baseline definition should be clearly defined and documented. The precision of the replicate initial analyses should be better than a 15% coefficient of variation (CV). Failure to obtain results within these parameters may indicate that the stability samples were prepared incorrectly and that consideration should be given to the preparation of another batch of samples. It may be an indication of analyte instability if a second batch of samples fails to meet these criteria; in such cases, methods to improve stability such as storage at lower temperatures or the addition of stabilizers to the samples should be considered.

Once a successful initial analysis result is obtained, the stability samples are assayed periodically as described above. Stability is indicated for chromatographic assays if the periodic analysis results are within 15% of nominal concentrations. Inherent method variability should be taken into account during long-term stability assessment, and a priori stability acceptance criteria may be defined based upon the observed performance of the method during validation. In general, for ligand binding assays it is recommended that the observed mean bias from the nominal concentration (or defined baseline) be ±20% (25% at the LLOQ and ULOQ).[4] Analysis at time point t_n may indicate that analyte degradation has occurred and stability is unacceptable, whereas

time point t_{n+1} may reveal that stability remains acceptable. Hence, it is recommended that 2 consecutive failing assessments, conducted on 2 successive days, be obtained before the maximum period for long-term stability is declared to have been reached.

Once long-term storage stability of the analyte in the matrix is defined, the need to prepare fresh matrix standards for each analytical run may be eliminated. Assay standards may be prepared in bulk, aliquoted, stored appropriately, and assigned an expiration date.

Short-Term Analyte Stability

Short-term stability, also referred to as process or bench-top stability, is evaluated to confirm that analyte degradation does not occur during the preparation/extraction of study samples prior to their analysis. To assess short-term stability, replicate stability samples at each concentration, prepared as described above, are removed from frozen storage and allowed to remain on the bench-top for the period of time for which stability is to be assessed (typically 4-24 hours). At the end of this period, an additional set of stability samples is removed from the freezer. Once the second set of samples has thawed, a set of matrix standards is prepared and both sets of stability samples are analyzed against the matrix standards. Stability is indicated for chromatographic assays if the difference in the analyzed results for the 2 sets of samples (ie, those maintained on the bench-top for the assessment period and those extracted immediately after thawing) is under 15% and the quantitated results are within 15% of the nominal values.

In specific cases it may have been determined during assay development that, for stability purposes, samples should be thawed in the refrigerator or on wet ice as opposed to on the bench-top. For such cases, short-term stability should be assessed under the conditions that will be used to thaw samples prior to analysis.

Freeze-Thaw Stability

From a practical standpoint, it is often necessary to subject samples to multiple freeze-thaw cycles before reportable analytical results may be obtained. Reasons necessitating analysis after multiple freeze-thaw cycles include failed analytical runs or the use of incorrect dilution factors (ie, initial results are outside the range of the standard curve).

Current guidance recommends the demonstration of freeze-thaw stability through at least 3 cycles.[1] To assess freeze-thaw stability, replicate (minimum n = 3) stability samples at low and high concentrations, prepared as described above, should be stored at their intended storage temperature for 24 hours and then thawed under the conditions specified in the assay protocol (bench-top, wet ice, warm water, or refrigerator). When completely thawed, the samples should be refrozen for at least 12 hours under the same conditions and thawed a second time. The freeze-thaw cycle should then be repeated a third time.

The samples that have been subjected to 3 freeze-thaw cycles are then analyzed together with a set of samples subjected to only 1 freeze-thaw cycle. To establish stability for chromatographic assays, the mean analyzed results for both sets of samples should be within 15% of the nominal sample concentrations. It is recommended that the precision (CV) for the replicate analyses not exceed 15%. A greater difference may indicate the potential for freeze-thaw instability if the number of freeze-thaw cycles is increased; an alteration of storage and/or thawing conditions should be considered in such cases.

Processed Sample Stability

Assessment of the ability to reanalyze sample extracts, the on-instrument stability of extracts, and the overall stability of sample extracts is needed to ensure processed sample stability.

Sample Extract Reanalysis/Reinjections Reproducibility

Sample extract reanalysis stability is typically assessed to determine whether it is possible to reinject/reanalyze processed samples in the event that their initial analysis is interrupted because of, for instance, instrument failure. One method to assess extract reanalysis stability is to prepare and extract a set of matrix standards and stability samples. The samples are then subjected to an initial instrumental analysis. Following the initial analysis, the processed samples (matrix standards and stability samples) are allowed to remain on the instrument for the time period that assesses sample extract reanalysis stability (typically 24-72 hours). The samples are then reanalyzed. After reanalysis, the results for the reanalyzed stability samples are calculated using both the standard curve derived from the initial analysis of the standards as well as that derived from the reanalyzed standards. Lack of a difference between these 2 results indicates that processed samples may be reanalyzed after the assessment period without the need to reanalyze calibration standards. If a difference between these results is observed, the results from the initial analysis of the QC samples, as calculated from the initial analysis of the matrix standards, should be compared with the results of the reanalyzed QC samples as calculated from the reanalyzed matrix standards. Agreement of these results indicates that an entire run sequence (matrix standards, samples, and QC samples) should be reinjected in the event that an instrument failure occurs in the middle of a run. If neither of these conditions is met, it may be concluded that processed samples cannot

be reanalyzed and that samples must be reprocessed in the event of instrument failure.

On-Instrument Stability

Matrix standard samples, QC samples, and study samples are generally analyzed in a serial vs a parallel manner. Thus, standards are analyzed at different points in time when compared with the study samples. For this reason, it is necessary to assess the stability of processed samples in the instrument over the anticipated run time of sample batches. This assessment may be accomplished by comparing the results of QC samples analyzed at the end of the run with those analyzed at the beginning of the run.

Assuming that study samples are bracketed by QC samples during the course of their analysis, data are generated with each batch of samples to demonstrate on-instrument stability. Thus, explicit additional experiments to assess on-instrument stability may not be required if the batch size during routine sample analysis happens to exceed that evaluated during assay validation.

Extract Stability

Extract stability assessment refers to an experiment in which the stability of stored sample extracts is assessed by their analysis against freshly prepared matrix standard extracts. The consensus reached during the AAPS/FDA workshop indicated that evaluation of extract stability, as defined herein, is generally not considered to be a part of routine validation testing.[4] This would typically be the case when all samples for a run (standards, QC, and study samples) are processed together. In cases where it is anticipated that extracted samples will be stored for a period of time prior to the extraction of the standard samples that they will be analyzed against, extract stability should be demonstrated for the anticipated duration of storage.

Reagent Stability

Stock Standard Solution Stability

As noted above, study samples are typically quantitated against matrix standard samples that are prepared by spiking stock solutions of the analyte into the biological matrix of interest. The analyte stock solutions are generally prepared in aqueous buffers, organic solvents, or mixtures thereof. Stability of the analyte stock solutions should be assessed as part of assay validation.

The stability of analyte stock solutions is independent of that of the reference standard material from which they are prepared. Thus, stock solutions should not routinely be assigned an expiration date matching that of the reference standard. Stock solution stability data should be generated to justify the period over which the solutions will be used. Stock solution stability is demonstrated by preparing a fresh solution from the reference material and comparing the absolute response of the fresh solution with that of the stored solution. In that study sample results are directly influenced by analyte stock solutions, it is recommended that the acceptable difference between the absolute responses of fresh stock solutions and aged stock solutions be tighter (within 5%-7%) than that normally applied to bioanalytical results (ie, within 15%-20%).

For large molecules, it is desirable that stock solution stability be demonstrated using fresh stock solutions prepared from the same lot of reference standard that was used to prepare the aged stock solutions. Such a requirement is generally not necessary for well-characterized small molecules.

Finally, newer stock solutions within their established stability period should not be used to assess the stability of an older solution. Rather, a fresh solution prepared directly from reference material should be used to determine the stability of any older stock.

Stability of Other Assay Reagents

Assay reagents other than analyte stock solutions whose stability affects assay performance are considered critical reagents. Generally, for small-molecule chromatographic-based analyses, most reagents (buffers, extraction solvents, etc) are considered noncritical; they are typically prepared at a frequency that makes their explicit stability determination moot. That is, batches are exhausted well before degradation occurs. It is recommended that, for chromatographic analyses, matrix standards, QC samples, and study samples be prepared using the same batches of assay reagents. In such cases, results meeting run acceptance criteria would indicate that instability of noncritical reagents did not affect run results. On the other hand, multiple sequential failed runs may indicate that a reagent initially believed to be noncritical had degraded to the point that analysis results were affected. In the event that investigation of the failed runs leads to the conclusion that the failures were due to the degradation of an assay reagent, the time period prior to the failed runs may be used to establish a reagent expiration date. The analytical protocol should then be amended to indicate that a fresh batch of the critical reagent should be prepared regularly, as opposed to when the reagent supply is exhausted.

Examples of reagents used for chromatographic assays whose stability should be determined during assay validation include derivatization reagent solutions and solutions of enzymes (eg, glucuronidase) used during sample preparation. The results obtained from the analysis of standard and QC samples prepared using fresh and aged solutions of

these reagents should be compared with assessments of the stability of the reagents in question.

For ligand binding assays, critical reagents may include antibodies, conjugated antibodies, enzymatic moieties, and biological matrices. Noncritical reagents typically include laboratory buffers and acidification reagents. Critical reagents require verification at lot number changes to ensure that the replacement reagent does not significantly alter the method outcome compared with the original reagent. For ligand binding methods, the most notable reagent requiring qualification is conjugated antibodies employed as part of the enzyme-linked immunosorbent assay detection system. The degree of conjugation and activity of the conjugated enzyme may vary from lot to lot, requiring an alteration of the quantity of conjugate used in the method with each lot of reagent.

For ligand binding assays, the method of reagent qualification should be fully described by the assay protocol or the general laboratory standard operating procedures (SOPs) under which the assay is run. The method may vary for each reagent. Only 1 critical reagent in a method should be qualified at a time. For ligand binding methods, qualification typically involves performing assays on a split solid-phase format using the replacement reagent in parallel with the original reagent. A combination of standards, validation samples, or QC samples is used to assess the similarity in response variable between the original and replacement reagents. Acceptance criteria should be defined a priori in an SOP.

For ligand binding assays, there may be instances when the original reagent is unavailable. In such cases, alternate approaches to reagent qualification should be described in laboratory SOPs. One possibility for qualification is to compare the instrument response signals obtained with the replacement reagent with that of historical data collected with the original reagent. Documentation of reagent qualification should be maintained and archived.

In regulated laboratories, it is required that all reagents be labeled with the identity, concentration or titer, storage conditions, and expiration date. Reagents obtained from a commercial source should arrive with this information. Caution should be used in that some vendors provide storage and expiration conditions that differ for opened and unopened reagents. For internal reagents the storage conditions and expiration dating may need to be determined empirically.

Conclusion

This article is a general guide to the stability assessment procedures and experiments that should be performed during bioanalytical assay validation. Additional experiments, often mandated by the nature of the analyte, may be required to ensure complete assessment of all stability-related factors that may influence analytical results. In short, this guidance should not serve as a substitute for good scientific judgment, which must come into play during all bioanalytical assay development activities.

References

1. US Department of Health and Human Services, FDA, CDER, CVM. *Guidance for the Industry: Bioanalytical Method Validation.* Washington, DC: US Department of Health and Human Services, FDA, CDER and CVM; 2001.

2. Shah VP, Midha KK, Findlay JWA, et al. Bioanalytical method validation—a revisit with a decade of progress. *Pharm Res*. 2000;17:1551-1557.

3. DeSilva B, Smith W, Weiner R, et al. Recommendations for the bioanalytical method validation of ligand-binding assays to support pharmacokinetic assessments of macromolecules. *Pharm Res*. 2003;20:1885-1900.

4. Viswanathan CT, Bansal S, Booth B, et al. Workshop/conference report—quantitative bioanalytical methods validation and implementation: best practices for chromatographic and ligand binding assays. *AAPS J [serial online]*. 2007;9:E30-E42.

5. Yau M. Stability of samples, standards, QC's, and reagent: FDA perspective. Quantitative bioanalytical methods validation and implementation—best practices for chromatographic and ligand binding assays. Paper presented at the AAPS/FDA Workshop; May 2, 2006; Arlington, VA.